Journal of Medical Virology 60:275–283 (2000)

Cloning, Sequencing, and Expression of the Hepatitis E Virus (HEV) Nonstructural Open Reading Frame 1 (ORF1)

Israrul Haque Ansari, Santosh Kumar Nanda, Hemlata Durgapal, Shipra Agrawal,[1] Sujit Kumar Mohanty,[1] Dinesh Gupta,[1] Shahid Jameel,[2] and Subrat Kumar Panda[1]*

[1]*Department of Pathology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, India*
[2]*Virology Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi, India*

Hepatitis E virus (HEV) causes enterically transmitted epidemic and sporadic viral hepatitis affecting millions of people in the developing world. Different geographical isolates of HEV show a high degree of homology at the nucleotide and amino acid levels. The ~7.2 kb RNA genome has three open reading frames of which ORF1 is predicted to code for the viral nonstructural polyprotein. The expression, processing and properties of the nonstructural ORF1 polyprotein have not been reported so far. In this study, the complete HEV ORF1 was reconstructed from overlapping fragments amplified by polymerase chain reaction (PCR) of total RNA isolated from the bile fluid of a rhesus monkey experimentally infected with HEV isolate from an epidemic. The complete assembled ORF1 was sequenced using HEV specific primers. The ORF1 polyprotein was expressed in *E. coli*, in a cell free translation system and in HepG2 cells, and was characterized by western blotting and immunoprecipitation using acute phase patient serum as well as polyclonal antibodies raised against defined parts of the ORF1 polyprotein. The nonstructural polyprotein of HEV was expressed as a 186 kDa protein. No processing was observed into discrete units, either in-vitro based on a kinetic analysis, or in HepG2 cells based on immunoprecipitation. ***J. Med. Virol. 60:275–283, 2000.*** © 2000 Wiley-Liss, Inc.

KEY WORDS: Hepatitis E virus (HEV), Polymerase Chain Reaction (PCR), Western blotting and Immunoprecipitation

INTRODUCTION

Hepatitis E virus (HEV) is a major causative agent of waterborne hepatitis and frequent epidemics attributed to this agent have been reported from different parts of the world [Panda et al., 1989; Khuroo et al., 1983; Reyes et al., 1990; Bradley, 1990]. The first such well characterized HEV epidemic was reported from India in 1955 [Vishwanathan, 1957]. HEV is a 27–34 nm non-enveloped virus, which has been classified on an interim basis as the prototypic member of the group of Hepatitis E like viruses (8th Report of the International Committee on Taxonomy of Viruses, in press).

The HEV genome is a positive-stranded RNA of about 7.2 kb, organized into a 5′ putative nonstructural region and a 3′ structural region. The genome contains three open reading frames (ORFs). ORF1, which is the largest in size, starts 27 nucleotides downstream of the 5′ end and terminates at position 5079, resulting in a protein of 1683 amino acids. The second ORF (ORF2) begins 37 nucleotides downstream of ORF1 and codes for a protein of 660 amino acids. The third ORF (ORF3), is the smallest in size, encoding a protein of 123 amino acids and overlaps with ORF1 at its terminal base [Tam et al., 1991]. Nonstructural ORF1 is believed to code for a putative polyprotein which has different motifs, such as those for a viral methyltransferase, a papain-like cysteine protease, a RNA helicase and a RNA dependant RNA polymerase [Koonin et al., 1992]. However none of these putative functional regions have been characterized so far. The structural ORFs have been expressed in prokaryotic and eukaryotic systems and immunogenicity of the resulting proteins have been reported in a number of studies [He et

This work was performed at the Department of Pathology, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, India.

Grant sponsor: Department of Science and Technology, India (SKP).

*Correspondence to: Prof. S.K. Panda, M.D., Department of Pathology, AIIMS, Ansari Nagar, New Delhi 110 029, India.
E-mail: skpanda@medinst.ernet.in or pandask@hotmail.com

Accepted 15 June 1999

TABLE I. Primers and Profiles for PCR Amplification of the Non-Structural ORF1

Nucleotide position			Profile	
From	To	Nucleotide sequence (5′ to 3′)	°C (min)	No. of cycles
1	27	gat atc ctc gag AGG CAG ACC ACA TAT GTG GTC GAT GC	95 (1), 55 (1), 73 (1.5)	35
1193	1217	CGA CGG ATC CCC TTG GAT ATA GCC T		
1369	1393	GGG CCG ACT CGT CAA AAA CCA ACA C		
1184	1207	TCC GCA CCC AGG CTA TAT CCA AGG	95 (1), 55 (1), 73 (1)	27
1205	1229	AGG GGA TCC GTC GTC TGG AAC GGG A		
2633	2656	GGG AGC AAG TTT CCC GAT AAG CAG		
2346	2369	GAC GGC CCG GCA CCG CCG CCT GCT	95 (1), 65 (4)	32
3803	3826	CGG GAT CCG GGC AGG GTG GTA GAA		
4169	4192	TGA ACT TGT TAC AAT CTT TCT GGA		
3771	3794	GCT TGG CCA CAG ACC TGT CCC TGT	95 (1), 55 (1), 73 (1.5)	35
4429	4462	TAG CAC ACT CTA GAC CCA GAG AAA AGT TAT TCT G		
4453	4476	ACA CTC CTC CAT AAT AGC ACA CTC		
4727	4753	GGA ATT CAC AGC CGG CGA TCA GGA CAG		
4438	4468	TTT TCT CTG GGT CTA GAG TGT GCT ATT ATG G	95 (1), 55 (1), 73 (1.5)	35
5131	5154	GGG CGC ATG GTC GCG AAC CCA TGG		
5373	5395	GAG TGG TCG GGC GGG TTG GCG AA		
4438	4468	TTT TCT CTG GGT CTA GAG TGT GCT ATT ATG G	95 (1), 55 (1), 73 (1.5)	35
5285	5307	GGG ATT GCG AAG GGC TGA GAA TC		
5373	5395	GAG TGG TCG GGC GGG TTG GCG AA		

al., 1993; Li et al., 1994; Panda et al., 1995]. We have earlier expressed ORF2 and ORF3 in animal cells. The ORF2 protein (pORF2) is an 88 kDa glycoprotein that is expressed intracellularly as well as on the cell surface [Jameel et al., 1996]. The ORF3 protein (pORF3) is a 13.5 kDa phosphoprotein, which is phosphorylated by the cellular mitogen activated protein kinase and associates with the cytoskeleton [Zafrullah et al., 1997]. This report describes the cloning and sequencing of ORF1 and expression of the putative HEV nonstructural polyprotein (pORF1) in prokaryotic and eukaryotic systems.

MATERIALS AND METHODS

Amplification, Cloning and Sequencing of ORF1

The 5131 nucleotide long ORF1 along with the 5′ noncoding region was amplified in several subgenomic fragments by reverse transcription-polymerase chain reaction (RT-PCR). The primers (Table I) were designed based on the sequence of the Myanmar strain of HEV (Acc. no. M73218) using the OLIGO-4 software. RNA was extracted from the bile of a rhesus monkey infected with fecal inoculum from a single patient collected during an epidemic outbreak of hepatitis E at Hyderabad, India [Jameel et al., 1992]. RNA extraction was carried out with the guanidinium-acid phenol method [Chomczynski and Sacchi, 1987] and converted into cDNA with 10 units of AMV Reverse Transcriptase (Promega, Madison, WI) at 42°C for 60 minutes. Reverse transcription was coupled to the first PCR in 1X PCR reaction buffer, 200 μM dNTPs, 2.5 mM $MgCl_2$, 50 pmols of each primer and 2.5 units of Taq DNA Polymerase (Life Technologies, Bethesda, MD) in combination with *Pfu*DNA polymerase (Stratagene, Germany). The 100 μl mix for coupled reverse transcription and first PCR contained 10 μl of the RNA template (extracted from 100 μl of bile). The second round of PCR was performed in a hemi-nested manner, where one primer was kept constant and the other was internal to one of the first round primers. For all amplifications Taq DNA Polymerase was used in combination with PfuDNA polymerase to avoid mutations during amplification. The reaction profile for each primer set used in the amplification is shown in Table I.

Each amplified fragment was purified from agarose gel and cloned into either pCR-Script SK(+) (Stratagene) or in pGEM-T (Promega) cloning vector according to the manufacturer's instructions. Multiple clones of each fragment were sequenced in both directions using T3, T7, KS, SP6 or HEV specific internal primers and the Sequenase version 2.0 sequencing kit (Amersham Int., UK) according to the manufacturer's protocol.

The fragments spanning nucleotides 1-1205, 159-862, 1288-1862, 1288-2540, 1205-2546, 2827-3667, 3093-3803, 4438-5131 and 3493-5154 were sequenced using multiple clones. Each of the sequenced fragments was expressed either in *E.coli* using the pRSET vector (Invitrogen, San Diego, CA), or in a rabbit reticulocyte coupled transcription and translation system using pSGI vector (Amersham Int., UK). It is a modified form of pSG5 eukaryotic expression vector (Stratagene) where a synthesized polylinker containing multiple cloning sites were inserted at EcoRI and BamHI sites [Jameel et al., 1996]. The expressed proteins were confirmed by immunoblot analysis or immunoprecipitation with acute phase HEV infected patient serum. Only those fragments that had confirmed HEV sequence and produced the expected immunoreactive proteins were used for reconstruction of full-length ORF1. For ease of assembly a unique BamHI site was created at nucleotide position 1208 in two of the amplification primers (forward and reverse). This led to a Met→Ile change at amino acid position 395 in ORF1.

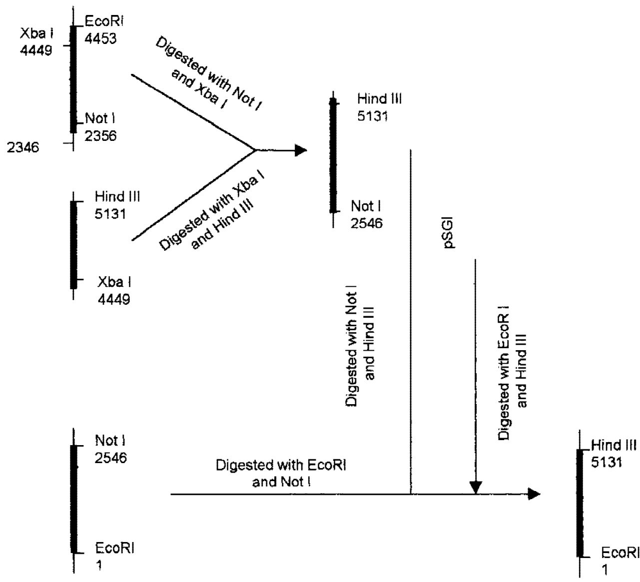

Fig. 1. Schematic representation of the strategy used for assembly of HEV ORF1. The unique sites used for the reconstruction are also mentioned.

The unique restriction sites present naturally within the ORF1 sequence were used for reconstruction of the full-length ORF1 from individual fragments. The complete reconstructed ORF1, in pRSET was sequenced again using internal HEV specific primers and the sequence deposited in data bank (Acc. no. AF028091). The insert spanning from 1-5131 nucleotides of HEV was taken out with XhoI and HindIII. The XhoI end was polished with the Klenow fragment of DNA polymerase I (Life Technologies) and cloned into plasmid pSGI digested with EcoRV and HindIII. A schematic representation of ORF1 reconstruction is presented in Figure 1. Complete details of subcloning and ORF1 reconstruction will be provided upon request.

Expression of the ORF1 in *E. coli*

The complete ORF1 was excised from the pSGI background as an EcoRI/HindIII fragment and cloned into pGEX-4T-2 (Pharmacia, Uppsala, Sweden) at the EcoRI and NotI sites within the multiple cloning sites. Before ligation HindIII site of insert and NotI site of vector was polished with Klenow fragment of DNA polymerase. Expression was carried out in *E.coli* JM109. A 5 ml overnight culture in LB broth was diluted into 500 ml of fresh medium and grown at 37°C with shaking till an A_{600} of 0.6 was achieved. Expression was induced with 1.0 mM IPTG (Isopropyl β-D-thiogalactoside) and the culture grown for another 3 hours. Expression of the Glutathione S-transferase (GST) fusion protein was evaluated on 6%–15% gradient SDS polyacrylamide gels stained with Coommassie Brilliant Blue (Sigma, St Louis, MO). Western blot was performed with *E.coli* expressed protein using either anti-GST antibodies (Pharmacia, Uppsala, Sweden) or with different polyclonal antibodies raised in rabbits against the predicted methyl transferase, helicase and RdRp regions.

Preparation of Anti-ORF1 Antibodies

Hexahistidine tagged fusion proteins were expressed in the prokaryotic expression vector pRSET (Invitrogen, CA) from three different regions of ORF1: nucleotides 159–862 within the methyltransferase region, nucleotides 3093–3803 within the helicase region, and nucleotides 4438–5131 within the RdRp region. The fusion proteins were purified on nickel NTA-agarose (Qiagen) as per manufacturer's guidelines. The reactivity of these proteins were confirmed by western blot assay using HEV infected patient serum. The rabbits were immunized with these purified proteins to raise antisera. Briefly, 500 μg of the purified recombinant protein was emulsified with an equal volume of Freund's complete adjuvant (DIFCO, Detroit, MI) and injected intradermally into rabbits at multiple sites. Three booster doses of the same amount of protein emulsified with incomplete Freund's adjuvant were given at 20-day intervals. Prior to the first immunisation, preimmune serum was collected and tested on western blots. Sera were collected aseptically from the immunised animals 30 days after the last booster, aliquoted and stored at –70°C for further use.

In Vitro Coupled Transcription and Translation

A coupled transcription-translation system (TNT, Amersham Int.) with bacteriophage T7 RNA polymerase was used for in vitro synthesis of the ORF1 polypeptide from plasmid pSGI-ORF1 according to the manufacturer's protocol. The synthesized protein was labelled with ^{35}S methionine-cysteine (promix, Amersham Intl) and a portion of the mixture was analysed by 6–15% SDS-polyacrylamide gel electrophoresis. The gel was treated with 0.5 M sodium salicylate (BDH, India) for 30 minutes at room temperature, dried and autoradiographed using Kodak X-Omat AR film.

Transfection and Metabolic Labelling of Cells

At ~50% confluency, HepG2 cells were transfected with plasmid pSG-ORF1 using Lipofectamine (Life Technologies) according to the manufacturer's guidelines. Two micrograms of supercoiled DNA was used for every 60mm tissue culture dish. Same amount of the vector (pSGI) served as a negative control. Forty four hours post-transfection, the cells were washed with phosphate-buffered saline (PBS pH 7.2) and incubated in 3 ml of methionine-cysteine deficient medium (Life Technologies) for 1 hr at 37°C in a CO_2 incubator. The cells were then labelled at 37°C for 4–6 hours with 150μCi ^{35}S methionine-cysteine in 1 ml of deficient medium. The labelled cells were washed twice with cold PBS pH 7.2 and lysed in 750 μl of RIPA buffer (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 5 mM Iodoacetamide,

0.5% Triton X-100, 1% SDS, 1% sodium deoxycholate, 2 mM phenylmethylsulfonyl fluoride). The lysate was cleared of cell debris by centrifuging at 12,000 rpm in a 220.87 VO1 rotor using refrigerated microfuge (Hermle Z323K, Germany) for 15 min at 4°C. The supernatant was collected and used for immunoprecipitation.

Immunoprecipitation

Ten microliters of the TNT mix in 500 μl of RIPA buffer or 750 μl of the clarified HepG2 cell lysate was incubated with 10 μl of the appropriate antiserum on ice for 1 hour. To this, 100 μl of a 10% suspension of Protein A Sepharose-4B (Pharmacia) was added, and the mixture was kept at 4°C with slow end to end shaking. After 1 hour, the reaction mixture was centrifuged for one minute at 10,000 rpm in a refrigerated microcentrifuge (Hermle Z323K, Germany). The supernatant was discarded and the beads were washed thrice with 1 ml of RIPA buffer, each time incubating at 4°C for 10 minutes with shaking. The washed beads were boiled with 50 μl of SDS-PAGE sample buffer and analysed on a 6%–15% gradient SDS polyacrylamide gel. The gel was treated with sodium salicylate, dried and exposed for autoradiography as described earlier.

Line Blot Analysis

The ORF1-GST fusion protein was purified from induced bacterial lysate by using commercial immunoaffinity purification system (Pharmacia). The purified ORF1-GST fusion protein along with purified ORF2 and ORF 3 proteins were vacuum blotted on to a nitrocellulose membrane (Schleicher and Schuell, Germany) by using a Hybri-Slot manifold (Life Technologies). Five hundred nanograms of each protein were blotted into individual slots respectively. The membrane was cut in strips of three slots for the individual HEV proteins and the strips were blocked in 5% skimmed milk powder in PBS [pH 7.2]. The strips were washed with wash buffer (0.3 M sodium chloride 0.05 M Phosphate buffer pH 7.2 with 0.1% Tween-20) and incubated with 1:100 dilution of sera, for one hour at 37°C. Thoroughly washed strips were incubated further with anti-human IgM-HRPO conjugate (Sigma) at 1:2000 dilution for one hour. Colour was developed with diaminobenzidine (Sigma0 in presence of hydrogen peroxide. Sera from 10 confirmed case of acute hepatitis E and 10 normal healthy individuals were analyzed for IgM antibodies.

RESULTS

Cloning and Sequencing of ORF1

The nonstructural ORF1 of HEV was amplified as multiple fragments by RT-PCR, sequence confirmed and expressed. These defined fragments were used for reconstruction of ORF1 using overlap PCR as well as unique restriction sites in the viral genome. Initially, ORF1 was reconstructed in two major parts, nucleotides 1–2546 and nucleotides 2666–5131. When expressed using an in vitro coupled transcription and translation system, these fragments produced polypeptides of the expected size (Fig. 2). Several smaller proteins of equal intensity were observed when the 2666–5131 fragment was expressed. These smaller polypeptides matched to the expected protein sizes produced from internal start codons.

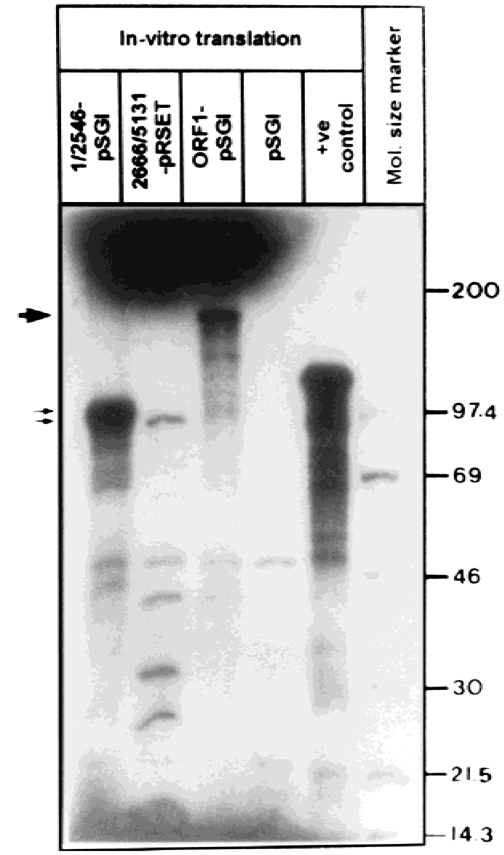

Fig. 2. Autoradiograph of the 6%–15% gradient SDS-PAGE showing labelled in-vitro coupled transcription and translation product of three independent clones from HEV covering nucleotides 1–2546-pSGI, 2666–5131-pRSET and 1–5131 in pSGI vector. The internal initiations in the insert 2666–5131 produced equally intense smaller size products, which correlates with the expected size. The vector pSGI andβ-galactosidase (~118kDa) were translated as negative and positive controls respectively. The ^{14}C labelled molecular size markers (in kilodaltons) are indicated (Amersham Int., UK).

Nucleotide sequence analysis was carried out after every reconstruction step and finally for the complete reconstructed fragment. The nonstructural region of the Indian strain of HEV showed 92%, 91%, 96%, and 77% homology at the nucleotide level with HEV strains from China (L08816), Pakistan (M80581), Myanmar (M73218) and Mexico (M74506), respectively. The amino acid homology was 96% to the Myanmar strain and 82% to the Mexican strain. Comparative homology estimates are presented in Table II. Changes in nucleotide sequence at the codon level were 0.67% for the first position, 0.35% for the second position and 2.2% for the third position. The 5′ nontranslated region (5′ NTR) of the Indian epidemic strain reported here was 27 nucleotides as designed by the primer in accordance with the Burmese strain. In the Mexican isolate, it is only 3 nucleotides long. A sporadic HEV isolate from

TABLE II. The Nucleotide and Amino Acid Sequence Homology of ORF1 Among Various Isolates of HEV

→ Nucleotide

	India AF028091	Myanmar M73218	Pakistan M80581	China L08816	Mexico M74506
India	—	96%	91%	92%	77%
Myanmar	96%	—	99%	93%	77%
Pakistan	96%	99%	—	94%	77%
China	95%	97%	98%	—	77%
Mexico	82%	83%	82%	83%	—

Amino Acid →

India (X98292) is reported to have a 24 nucleotide long 5′ NTR with A to C substitution at position 11. This falls within the proposed stem-loop structure generated at the 5′ end of the genome [Huang et al., 1992].

Upon secondary structure analysis the putative RNA dependent RNA polymerase (RdRp) region showed a unique RNA binding domain [Siomi and Dreyfuss, 1997] of "βαββαβ" between amino acid positions 1567–1652 with the following motifs: AVLI, LKLKVDFR, MRLYAGVVV, LPDVVRFAG, ERADELRIAVSDFL-RKLTNVA and QMCVDVVSRVYGV. These motifs were present in all other reported HEV sequences and the particular polypeptide stretch was found to bind at the 3′ nontranslated region of HEV RNA [Panda et al., unpublished data]. This would be critical for the replication of plus-stranded HEV genomic RNA into the minus-stranded replication intermediate.

Several other conserved motifs [Koonin et al., 1992] were also found in the deduced amino acid sequence of ORF1. Six conserved motifs for viral methyltransferases, identified between amino acid positions 71 to 249, were WNQPIQRV, SGRCLEIGAH, PNVVHRCF, RD-VQRWYTAPT, CFDGFSGCSCP and SAGYNHDVSN-LRSWI. Seven consensus domains of RNA helicases were identified between amino acid positions 968 to 1184 as GCRVTPGVVQYQFTAGVPGSGKS, VVVVP-TRELR, GRRVVIDEAP, HLLGDPNQI, VTHRCP, PVHDSQGATYTYTTI and VALTRHTEKW. Two well defined conserved regions, GVPGSGKS and DEAP were also found in the putative NTPase region of the nonstructural protein between amino acid positions 975-1029. The hypothetical active site of papain-like cysteine proteases, with cysteine and histidine residues at conserved positions [Koonin et al., 1992] were also seen in this strain.

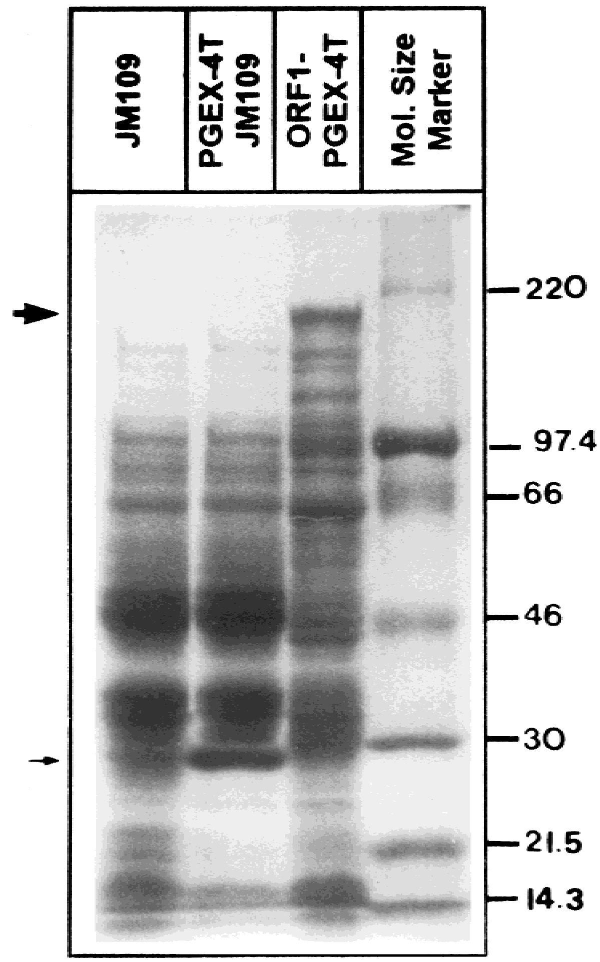

Fig. 3. *E.coli* expression of the nonstructural ORF1 of HEV in prokaryotic system. The *E.coli* (JM1O9), *E.coli* transformed with parent vector (pGEX-4T) and ORF1-pGEX-4T-2 were induced with 1.0 mM IPTG and analyzed on 6%–15% gradient SDS-PAGE followed by staining with Coommassie Brilliant Blue. The big arrow represents the induced ORF1 protein while small arrow indicates expressed glutathione S transferase protein. Molecular size markers (in kilodaltons) are also indicated (Rainbow markers, Amersham Int., UK).

Prokaryotic Expression of pORF1

The nonstructural ORF1 and 5′NCR, spanning nucleotides 1 to 5131 was reconstructed and cloned in the prokaryotic expression vector pGEX-4T-2 as described above. On induction with IPTG, a fusion protein of the expected size of ~212 kDa was expressed as observed by analysis of the total *E.coli* lysate on a 6%–15% gradient SDS polyacrylamide gel (Fig. 3). This polypeptide was absent from control *E.coli* (JM109) as well as cells transformed with the parent vector plasmid. On western blot analysis also, strong reactivity of this polypeptide was observed with antibodies directed against the helicase and RdRp regions of the HEV nonstructural polyprotein (Fig. 4). Specific, though weaker reactivity was also observed on western blots with antibodies to the methyltransferase region or the Glutathione-S-transferase fusion partner. Several smaller bands were also identified on western blots and may

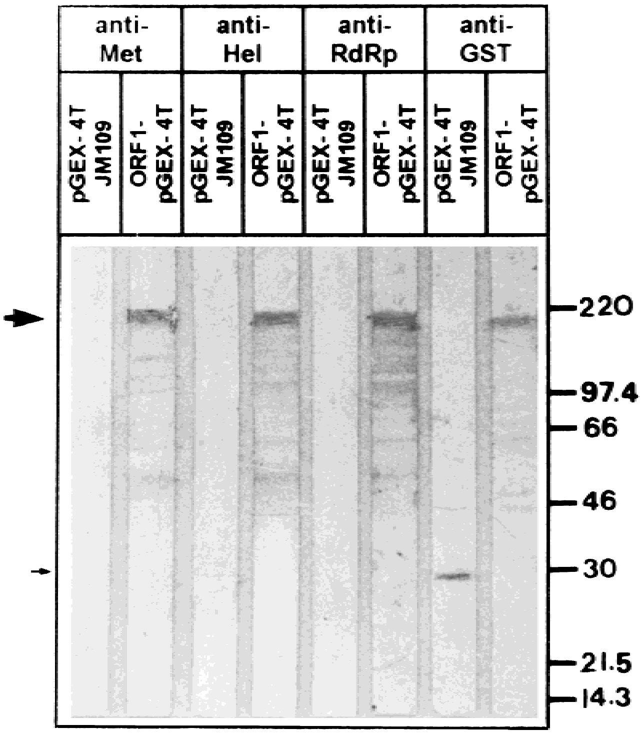

Fig. 4. Western blot analysis of *E.coli* expressed ORF1. The parent vector (pGEX-4T) along with ORF1-pGEX-4T was induced and resolved on 6%–15% gradient SDS-PAGE. Immunostaining was carried out with rabbit anti-Met, anti-Hel, anti-RdRp and anti-GST antibodies. GST was identified at ~26 kDa (small arrow) and recombinant protein of ~212 kDa (big arrow) along with other smaller possible degradation products. Positions of molecular size markers (in kilodaltons) are also indicated.

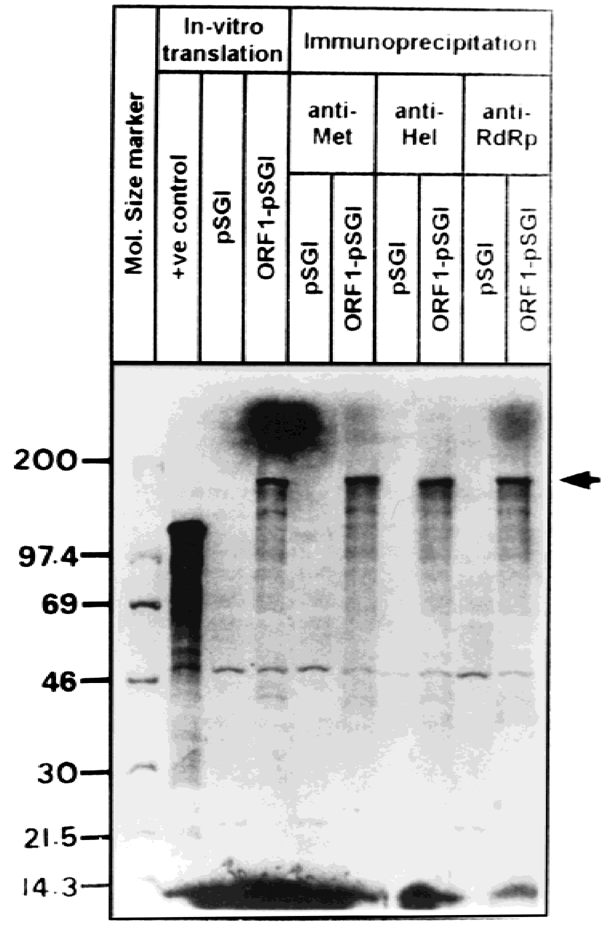

Fig. 5. Autoradiograph showing in-vitro coupled transcription and translation of ORF1-pSGI and immunoprecipitation of labelled protein. Immunoprecipitation was carried out with rabbit anti-Met, anti-Hel and anti-RdRp antibodies. Mock experiment with pSGI was carried out in the same manner with all the three antibodies independently, to serve as control. The immunoprecipitates were analyzed on 6%–15% SDS-PAGE and visualized by flourography. βgalactosidase (~118kDa) was used as a positive control for in-vitro translation. The molecular sizes of ^{14}C labelled markers (in kilodaltons) are also indicated (Amersham Int., UK).

represent degradation products of the ~212 kDa GST-ORF1 fusion protein.

Eukaryotic Expression of pORF1

The HEV ORF1, cloned in vector pSGI containing the SV40 early promoter-enhancer, was used to express the nonstructural polypeptide in eukaryotic systems. Initially the expression of ORF1 was carried out in vitro using coupled transcription and translation (TNT) system. A polypeptide of ~186 kDa, well in agreement with the expected size of ORF1 was expressed and immunoprecipitated with all the three polyclonal antibodies directed to the methyltransferase, helicase and RdRp regions of the ORF1 polyprotein (Fig. 5). In-vitro coupled transcription and translation was performed at 30°C for 2 hours and the product was incubated at 37°C for various lengths of time. No processing was observed even up to 24 hrs (data not shown). For expression in animal cells, HepG2 cells were transfected with the expression vector pSG-ORF1, pulse labelled with promix and immunoprecipitated with polyclonal antibodies directed against the RdRp region. A protein of the expected size was immunoprecipitated from pSG-ORF1 transfected cells, but not from the control vector transfected cells (Fig. 6). Other nonspecific cellular products were also observed, as they were also present in control cells.

Diagnostic Potential of the Nonstructural Polyprotein

Antibodies of IgM class were detected against ORF1 in 10 samples from patients with acute HEV infection tested by line immunoblot assay (Fig. 7). Similarly IgM class antibodies were also detected against the *E.coli* expressed and purified ORF2 and ORF-3 proteins in these sera. However, all the 10 sera from healthy individuals were negative for IgM class antibodies.

DISCUSSION

Infection due to HEV accounts for about a third of the sporadic acute viral hepatitis observed in the Indian subcontinent [Khuroo et al., 1983; Panda et al., 1989], and is a major health problem in subtropical and tropical parts of the world. It is important to understand the biology of this virus in order to develop appropriate control measures. The genome of HEV has

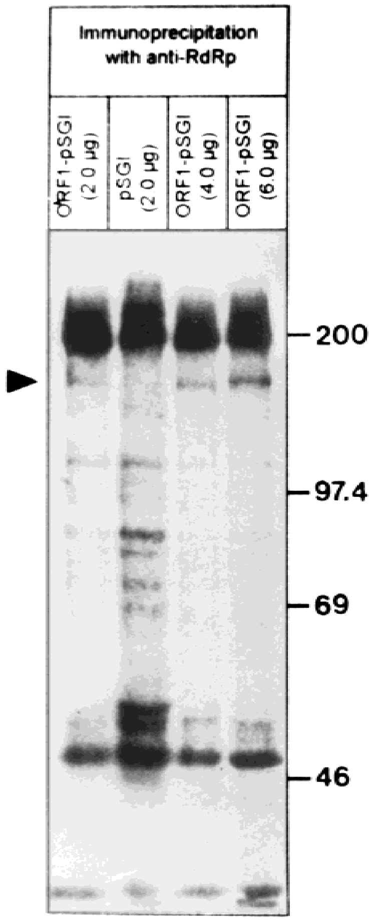

Fig. 6. Immunoprecipitation of in-vivo translated ORF1-pSGI with anti-RdRp. HepG2 cells were transfected with 2, 4, and 6.0 μg of double CsCl purified ORF1-pSG1 plasmid. 2.0 μg of the parent vector pSGI served as control. The transfected cells were pulse labelled with promix and immunoprecipitates were analyzed on 8% SDS-PAGE and visualized by flourography. Positions of the molecular size markers and ORF1 protein is indicated.

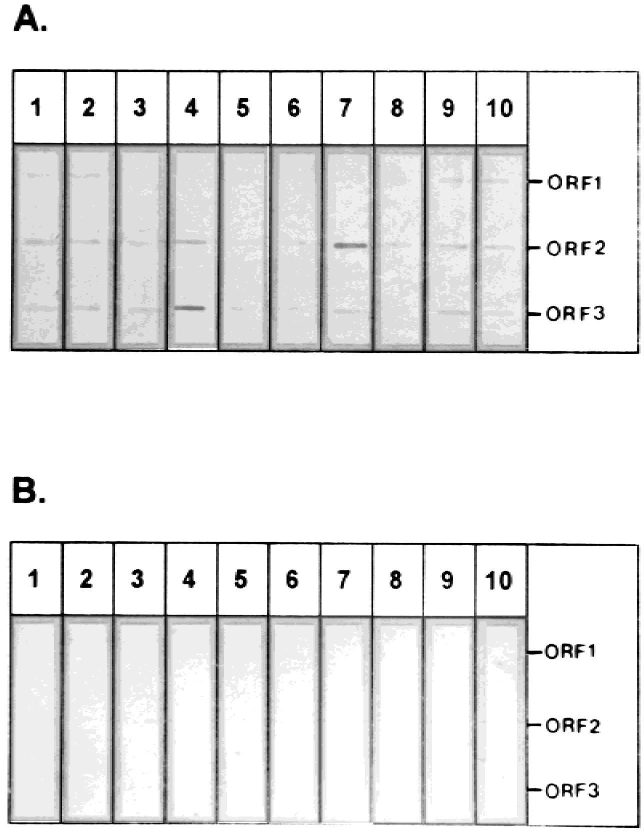

Fig. 7. Line immunoassay showing IgM anti-HEV posititvity against *E.coli* expressed and purified ORF1, ORF2 and ORF3 proteins. Panel A = 10 acute HEV infected patients' sera. Panel B = 10 normal control sera.

been cloned and sequenced from several different geographical isolates [Tam et al., 1991; Aye et al., 1992; Huang et al., 1992; Tsarev et al., 1992]. However, the viral biology is poorly understood due to the lack of a suitable cell culture system. The alternative approach is to express the viral proteins in vitro, study their properties and thereby define their role in the viral life cycle. Our earlier studies involved such a subgenomic expression strategy to characterize the HEV structural proteins, pORF2 and pORF3 [Panda et al., 1995; Jameel et al., 1996; Zafrullah et al., 1997; Zafrullah et al., 1999]. In this study, a similar strategy was used to gain insight into the nonstructural protein (pORF1) of HEV.

The nonstructural part of HEV genome consists of a single large open reading frame with putative domains for enzymatic functions necessary for viral RNA replication and protein processing. The HEV ORF1 was amplified and cloned in several fragments. Each fragment was sequenced, expressed in *E.coli* and its immunogenicity was checked before being used for reconstruction of the entire ORF1. The reconstructed ORF1 was then expressed in *E.coli*, in HepG2 cells and in an in vitro coupled transcription and translation system.

The full length ORF1 was expressed in the eukaryotic system both in transfected cells as well as in vitro translation. The product was of ~186 kDa, well in agreement with the predicted size of the ORF1. Cell free translation system uses the first AUG as the first codon of the protein to be synthesised. However in the case of HEV there are two more AUG codons (nucleotide positions 14 and 23), which appear before the ORF1 AUG situated at nucleotide position 27. Any protein that is synthesized from the AUG codons at nucleotide position 14 or 23 would be very short, terminating at nucleotide position 89. Therefore the protein expressed in cell free translation system is likely to span nucleotides from 27 to 5083. Immunoprecipitation of the radiolabelled protein produced in cell free translation system gave a single band and did not show any self processing even after 24 hrs of incubation at 37°C (data not shown). In the transfection studies in HepG2 cells a protein of the expected size (~186 kDa) was immunoreactive with anti-RdRp antibody. The cell free translated labelled protein did not show any processivity upon incubation with 100 mM of different metal ions i.e. Mg^{++}, Ca^{++}, Co^{++}, Mn^{++}, Zn^{++}, and K^{+} independently for 6 hours. The cell free translated ORF1

was also incubated with separately translated unlabelled protease domain to look for any trans cleavage of the protein. However the protein was intact and no signal for any discrete/processed domain/unit was found even after overnight incubation (data not shown). Interestingly in an infectious cDNA clone produced using this ORF1 fragment, complete processing of the ORF1 protein was observed and the proteins corresponding to methyl transferase, helicase and RNA dependent RAN polymerase (RdRp) domains were identified following immunoprecipitation with specific antibodies [Panda et al. under review]. In this context it is possible that the processing of the polyprotein occurs only when the other viral proteins are present. The possible role of ORF3 protein that is phosphorylated by MAP kinase in activating the viral/cellular protease to initiate processing cannot be ruled out [Zafrullah et al., 1997].

The prokaryotic expression was more difficult because of the large size of the protein and possible problems of codon utilization. The smaller subgenomic fragments (159–862, 1205–2546, 2827–3667, 3093–3803, 4438–5131 and 3493–5154) were expressed individually using the pRSET vector system (data not shown). However, our attempts to express the complete protein in the same vector as well as a few other vectors (pQE; Qiagen, pT7; Pharmacia) were not successful. Finally, the expression was achieved in pGEX-4T vector system which produces a GST fusion protein. In this system the complete fusion protein of about ~212 kDa was produced as expected. On immunostaining of the western blot with anti-GST antibody several smaller products were detected. It is strongly suspected that these might be due to degradation of ORF1, as they were detected in immunoblot analysis not only by anti-GST but also with other antibodies (anti-Met, anti-Hel, anti-RdRp). Antibodies of IgM class were detected in all the cases of acute HEV hepatitis. These antibodies were absent in normal controls. This indicates an IgM class immune response during acute infection. Therefore it is possible to use in serodiagnosis of acute HEV infection. It is known that the antibody responsible for virus immunoagglutination disappears rapidly during convalescence [Mushahwar et al., 1993; Favorov et al., 1996]. Therefore, detailed informations regarding appearance and disappearance of IgM and IgG class antibodies against all the three viral proteins following HEV infection are required for the development of reliable diagnostics. The situation is further complicated in the endemic zone where the exposure to HEV is perennial and IgG antibodies exists in a high proportion of normal individuals. The availability of all the three recombinant viral proteins could throw light on this aspect by doing time kinetic analysis.

The genome sequence was analysed for the reported putative domains of methyl transferase, helicase, and RNA dependent RNA polymerase. In addition, the putative RNA binding domain of RdRp was identified from the conserved motifs in secondary structure [Siomi and Dreyfuss, 1997]. This domain expressed as a subgenomic polypeptide binds to the 3′ noncoding region of the RNA sense strand [Panda, unpublished data]. In the context of whole viral genome transfection, processing of the ORF1 polyprotein to domain specific products have been identified [Panda, unpublished data].

The present report describes for the first time the expression of complete ORF1 protein of HEV and it's possible use in diagnosis. Further studies on the functional domains as well as processing can help in the development of antiviral capable of inhibiting viral growth and possibly the severe consequences of infection.

REFERENCES

Aye TT, Uchida T, Ma XZ, Iida F, Shikata T, Zuang H , Win KM. 1992. Complete nucleotide sequence of a hepatitis E virus isolated from the Xinjiang epidemic. 1986–1988) of China. Nucleic Acids Res 20:3512.

Bradley DW. 1990. Enterically transmitted non-A, non-B hepatitis. Br Med Bull 46: 442–461.

Chomczynski P, Sacchi N. 1987. Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159.

Favorov MO, Khudyakov YE, Mast EE, Yashina TL, Shapiro CN, Khudyakova NS, Jue DL, Onischenko GG, Margolis HS, Fields HA. 1996. IgM and IgG antibodies to hepatitis E virus HEV) detected by an enzyme immunoassay based on an HEV-specific artificial recombinant mosaic protein. J Med Virol 5:50–58.

He JA, Tam AW, Yarbough PO, Reyes GR, Carl M. 1993. Expression and diagnostic utility of hepatitis E virus putative structural proteins expressed in insect Cells. J Clin Microbiol 31:2167–73.

Huang, CC, Nguyen D, Fernandez J, Yun KY, Fry KE, Bradley DW, Tam AW, Reyes GR. 1992. Molecular cloning and sequencing of the Mexico isolate of hepatitis E virus (HEV). Virology 191:550–558.

Jameel S, Durgapal H, Habibullah CM, Khuroo MS, Panda SK. 1992. Enteric non-A, non-B hepatitis: epidemics, animal transmission, and hepatitis E virus detection by polymerase chain reaction. J Med Virol 37:263–270.

Jameel S, Zafrullah M, Özdener MH, Panda SK. 1996. Expression in animal cells and characterization of the hepatitis E virus structural proteins. J Virol 70:207–216.

Khuroo MS, Deurmeyer W, Zargar SA, Ahanger MA, Shah MA. 1983. Acute sporadic non-A, non-B hepatitis in India. Am J of Epidemiol 118:360–364.

Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. 1992. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of new group of animal and plant positive-strand RNA viruses. Proc Nat Acad Sci USA 89:8259–63.

Li F, Zhuang H, Kolivas S, Locarnini SA, Anderson DA. 1994. Persistent and transient antibody responses to hepatitis E virus detected by western immunoblot using open reading frame 2 and 3 and glutathione S-transferase fusion proteins. J Clin Microbiol 32: 2060–2066.

Mushahwar IK, Dawson GJ, Bile KM, Magnius LO. 1993. Serological studies of an enterically transmitted non-A, non-B hepatitis in Somalia. J Med Virol 40:218–221.

Panda SK, Datta R, Kaur K, Zuckerman AJ, Nayak NC. 1989. Enterically transmitted non-A, non-B hepatitis: recovery of virus-like particles from an epidemic in South Delhi and transmission studies in rhesus monkeys. Hepatology 10:466–472.

Panda SK, Nanda SK, Zafrullah M, Ansari IH, Özdener MH, Jameel S. 1995. An Indian strain of hepatitis E virus (HEV): cloning, sequence, and expression of structural region and antibody response in sera from individuals from an area of high-level HEV endemicity. J Clin Microbiol 33:2653–2659.

Reyes GR, Purdy MA, Kim J, Luk KC, Young LM, Fry KM, Bradley DW. 1990. Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non–B hepatitis. Science 247:1335–1339.

Siomi H, Dreyfuss G. 1997. RNA-binding proteins as regulator of gene expression. Current Opinion in Genetics and Development. 7:345–353.

Tam AW, Smith MM, Guerra ME, Huang CC, Bradley DW, Fry KE, Reyes GR. 1991. Hepatitis E virus (HEV): molecular cloning and sequencing of the full-length viral genome. Virology 185:120–31.

Tsarev S, Emerson SU, Reyes GR, Tsareva TS, Legters LJ, Malik IA, Iqbal M, Purcell RH. 1992. Characterization of a prototype Strain of hepatitis E virus. Proc Nat Acad of Sci USA 89:559–563.

Vishwanathan R. 1957. Infectious hepatitis in Delhi (1955–56). Indian J Med Res 45:Suppl 1–30.

Zafrullah M, Özdener MH, Kumar R, Panda SK, Jameel S. 1999. Mutational Analysis of Glycosylation, Membrane Translocation and Cell Surface Expression of the Hepatitis E Virus ORF2 Protein. J Virol 73(5):4074–82.

Zafrullah M, Özdener MH, Panda SK, Jameel S. 1997. The ORF3 protein of HEV is a phosphoprotein that associates with cytoskeleton. J Virol 71:9045–9053.